RAW SEQUENCE LISTING EFS

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 10/663,857A
Source: IFW16-EFS
Date Processed by STIC: 11/9/06

ENTERED

IFW16

RAW SEQUENCE LISTING DATE: 11/09/2006
PATENT APPLICATION: **US/10/663,857A** TIME: 13:06:10

Input Set : **N:\efs\10663857a_efs\PX018432US_SEQ_1stOA_Amended.txt**
Output Set: **N:\CRF4\11092006\J663857A.raw**

```
      1 <110> APPLICANT: Samsung Electronics Co. Ltd
      3 <120> TITLE OF INVENTION: A variant of HNF-1a gene having novel single nucleotide
      4       polymorphism and a variant protein encoded by the same
      6 <130> FILE REFERENCE: YPL-0064
C-->  8 <140> CURRENT APPLICATION NUMBER: US/10/663,857A
      9 <141> CURRENT FILING DATE: 2003-09-15
     11 <160> NUMBER OF SEQ ID NOS: 27
     13 <170> SOFTWARE: KopatentIn 1.71
     15 <210> SEQ ID NO: 1
     16 <211> LENGTH: 1896
     17 <212> TYPE: DNA
     18 <213> ORGANISM: Homo sapiens
     20 <220> FEATURE:
     21 <221> NAME/KEY: CDS
     22 <222> LOCATION: (1)..(1893)
     23 <223> OTHER INFORMATION: amino acid sequence of HNF-1a
     26 <400> SEQUENCE: 1
     27 atg gtt tct aaa ctg agc cag ctg cag acg gag ctc ctg gcg gcc ctg          48
     28 Met Val Ser Lys Leu Ser Gln Leu Gln Thr Glu Leu Leu Ala Ala Leu
     29   1               5                  10                  15
     31 ctc gag tca ggg ctg agc aaa gag gca ctg atc cag gca ctg ggt gag          96
     32 Leu Glu Ser Gly Leu Ser Lys Glu Ala Leu Ile Gln Ala Leu Gly Glu
     33              20                  25                  30
     35 ccg ggg ccc tac ctc ctg gct gga gaa ggc ccc ctg gac aag ggg gag         144
     36 Pro Gly Pro Tyr Leu Leu Ala Gly Glu Gly Pro Leu Asp Lys Gly Glu
     37          35                  40                  45
     39 tcc tgc ggc ggc ggt cga ggg gag ctg gct gag ctg ccc aat ggg ctg         192
     40 Ser Cys Gly Gly Gly Arg Gly Glu Leu Ala Glu Leu Pro Asn Gly Leu
     41      50                  55                  60
     43 ggg gag act cgg ggc tcc gag gac gag acg gac gac gat ggg gaa gac         240
     44 Gly Glu Thr Arg Gly Ser Glu Asp Glu Thr Asp Asp Asp Gly Glu Asp
     45  65                  70                  75                  80
     47 ttc acg cca ccc atc ctc aaa gag ctg gag aac ctc agc cct gag gag         288
     48 Phe Thr Pro Pro Ile Leu Lys Glu Leu Glu Asn Leu Ser Pro Glu Glu
     49                  85                  90                  95
     51 gcg gcc cac cag aaa gcc gtg gtg gag acc ctt ctg cag gag gac ccg         336
     52 Ala Ala His Gln Lys Ala Val Val Glu Thr Leu Leu Gln Glu Asp Pro
     53              100                 105                 110
     55 tgg cgt gtg gcg aag atg gtc aag tcc tac ctg cag cag cac aac atc         384
     56 Trp Arg Val Ala Lys Met Val Lys Ser Tyr Leu Gln Gln His Asn Ile
     57          115                 120                 125
     59 cca cag cgg gag gtg gtc gat acc act ggc ctc aac cag tcc cac ctg         432
     60 Pro Gln Arg Glu Val Val Asp Thr Thr Gly Leu Asn Gln Ser His Leu
```

RAW SEQUENCE LISTING DATE: 11/09/2006
PATENT APPLICATION: **US/10/663,857A** TIME: 13:06:10

Input Set : **N:\efs\10663857a_efs\PX018432US_SEQ_1stOA_Amended.txt**
Output Set: **N:\CRF4\11092006\J663857A.raw**

```
61     130                 135                 140
63 tcc caa cac ctc aac aag ggc act ccc atg aag acg cag aag cgg gcc          480
64 Ser Gln His Leu Asn Lys Gly Thr Pro Met Lys Thr Gln Lys Arg Ala
65 145                 150                 155                 160
67 gcc ctg tac acc tgg tac gtc cgc aag cag cga gag gtg gcg cag cag          528
68 Ala Leu Tyr Thr Trp Tyr Val Arg Lys Gln Arg Glu Val Ala Gln Gln
69                 165                 170                 175
71 ttc acc cat gca ggg cag gga ggg ctg att gaa gag ccc aca ggt gat          576
72 Phe Thr His Ala Gly Gln Gly Gly Leu Ile Glu Glu Pro Thr Gly Asp
73             180                 185                 190
75 gag cta cca acc aag aag ggg cgg agg aac cgt ttc aag tgg ggc cca          624
76 Glu Leu Pro Thr Lys Lys Gly Arg Arg Asn Arg Phe Lys Trp Gly Pro
77         195                 200                 205
79 gca tcc cag cag atc ctg ttc cag gcc tat gag agg cag aag aac cct          672
80 Ala Ser Gln Gln Ile Leu Phe Gln Ala Tyr Glu Arg Gln Lys Asn Pro
81     210                 215                 220
83 agc aag gag gag cga gag acg cta gtg gag gag tgc aat agg gcg gaa          720
84 Ser Lys Glu Glu Arg Glu Thr Leu Val Glu Glu Cys Asn Arg Ala Glu
85 225                 230                 235                 240
87 tgc atc cag aga ggg gtg tcc cca tca cag gca cag ggg ctg ggc tcc          768
88 Cys Ile Gln Arg Gly Val Ser Pro Ser Gln Ala Gln Gly Leu Gly Ser
89                 245                 250                 255
91 aac ctc gtc acg gag gtg cgt gtc tac aac tgg ttt gcc aac cgg cgc          816
92 Asn Leu Val Thr Glu Val Arg Val Tyr Asn Trp Phe Ala Asn Arg Arg
93             260                 265                 270
95 aaa gaa gaa gcc ttc cgg cac aag ctg gcc atg gac acg tac agc ggg          864
96 Lys Glu Glu Ala Phe Arg His Lys Leu Ala Met Asp Thr Tyr Ser Gly
97         275                 280                 285
99 ccc ccc cca ggg cca ggc ccg gga cct gcg ctg ccc gct cac agc tcc          912
100 Pro Pro Pro Gly Pro Gly Pro Gly Pro Ala Leu Pro Ala His Ser Ser
101     290                 295                 300
103 cct ggc ctg cct cca cct gcc ctc tcc ccc agt aag gtc cac ggt gtg          960
104 Pro Gly Leu Pro Pro Pro Ala Leu Ser Pro Ser Lys Val His Gly Val
105 305                 310                 315                 320
107 cgc tat gga cag cct gcg acc agt gag act gca gaa gta ccc tca agc         1008
108 Arg Tyr Gly Gln Pro Ala Thr Ser Glu Thr Ala Glu Val Pro Ser Ser
109                 325                 330                 335
111 agc ggc ggt ccc tta gtg aca gtg tct aca ccc ctc cac caa gtg tcc         1056
112 Ser Gly Gly Pro Leu Val Thr Val Ser Thr Pro Leu His Gln Val Ser
113             340                 345                 350
115 ccc acg ggc ctg gag ccc agc cac agc ctg ctg agt aca gaa gcc aag         1104
116 Pro Thr Gly Leu Glu Pro Ser His Ser Leu Leu Ser Thr Glu Ala Lys
117         355                 360                 365
119 ctg gtc tca gca gct ggg ggc ccc ctc ccc cct gtc agc acc ctg aca         1152
120 Leu Val Ser Ala Ala Gly Gly Pro Leu Pro Pro Val Ser Thr Leu Thr
121     370                 375                 380
123 gca ctg cac agc ttg gag cag aca tcc cca ggc ctc aac cag cag ccc         1200
124 Ala Leu His Ser Leu Glu Gln Thr Ser Pro Gly Leu Asn Gln Gln Pro
125 385                 390                 395                 400
```

RAW SEQUENCE LISTING DATE: 11/09/2006
PATENT APPLICATION: **US/10/663,857A** TIME: 13:06:10

Input Set : **N:\efs\10663857a_efs\PX018432US_SEQ_1stOA_Amended.txt**
Output Set: **N:\CRF4\11092006\J663857A.raw**

```
127 cag aac ctc atc atg gcc tca ctt cct ggg gtc atg acc atc ggg cct        1248
128 Gln Asn Leu Ile Met Ala Ser Leu Pro Gly Val Met Thr Ile Gly Pro
129                 405                 410                 415
131 ggt gag cct gcc tcc ctg ggt cct acg ttc acc aac aca ggt gcc tcc        1296
132 Gly Glu Pro Ala Ser Leu Gly Pro Thr Phe Thr Asn Thr Gly Ala Ser
133             420                 425                 430
135 acc ctg gtc atc ggc ctg gcc tcc acg cag gca cag agt gtg ccg gtc        1344
136 Thr Leu Val Ile Gly Leu Ala Ser Thr Gln Ala Gln Ser Val Pro Val
137         435                 440                 445
139 atc aac agc atg ggc agc agc ctg acc acc ctg cag ccc gtc cag ttc        1392
140 Ile Asn Ser Met Gly Ser Ser Leu Thr Thr Leu Gln Pro Val Gln Phe
141     450                 455                 460
143 tcc cag ccg ctg cac ccc tcc tac cag cag ccg ctc atg cca cct gtg        1440
144 Ser Gln Pro Leu His Pro Ser Tyr Gln Gln Pro Leu Met Pro Pro Val
145 465                 470                 475                 480
147 cag agc cat gtg acc cag aac ccc ttc atg gcc acc atg gct cag ctg        1488
148 Gln Ser His Val Thr Gln Asn Pro Phe Met Ala Thr Met Ala Gln Leu
149                 485                 490                 495
151 cag agc ccc cac gcc ctc tac agc cac aag ccc gag gtg gcc cag tac        1536
152 Gln Ser Pro His Ala Leu Tyr Ser His Lys Pro Glu Val Ala Gln Tyr
153             500                 505                 510
155 acc cac acg ggc ctg ctc ccg cag act atg ctc atc acc gac acc acc        1584
156 Thr His Thr Gly Leu Leu Pro Gln Thr Met Leu Ile Thr Asp Thr Thr
157         515                 520                 525
159 aac ctg agc gcc ctg gcc agc ctc acg ccc acc aag cag gtc ttc acc        1632
160 Asn Leu Ser Ala Leu Ala Ser Leu Thr Pro Thr Lys Gln Val Phe Thr
161     530                 535                 540
163 tca gac act gag gcc tcc agt gag tcc ggg ctt cac acg ccg gca tct        1680
164 Ser Asp Thr Glu Ala Ser Ser Glu Ser Gly Leu His Thr Pro Ala Ser
165 545                 550                 555                 560
167 cag gcc acc acc ctc cac atc ccc agc cag gac cct gcc ggc atc cag        1728
168 Gln Ala Thr Thr Leu His Ile Pro Ser Gln Asp Pro Ala Gly Ile Gln
169                 565                 570                 575
171 cac ctg cag ccg gcc cac cgg ctc agc gcc agc ccc aca gtg tcc tcc        1776
172 His Leu Gln Pro Ala His Arg Leu Ser Ala Ser Pro Thr Val Ser Ser
173             580                 585                 590
175 agc agc ctg gtg ctg tac cag agc tca gac tcc agc aat ggc cag agc        1824
176 Ser Ser Leu Val Leu Tyr Gln Ser Ser Asp Ser Ser Asn Gly Gln Ser
177         595                 600                 605
179 cac ctg ctg cca tcc aac cac agc gtc atc gag acc ttc atc tcc acc        1872
180 His Leu Leu Pro Ser Asn His Ser Val Ile Glu Thr Phe Ile Ser Thr
181     610                 615                 620
183 cag atg gcc tct tcc tcc cag         taa                           1896
184 Gln Met Ala Ser Ser Ser Gln
185 625                 630
188 <210> SEQ ID NO: 2
189 <211> LENGTH: 631
190 <212> TYPE: PRT
191 <213> ORGANISM: Homo sapiens
```

RAW SEQUENCE LISTING DATE: 11/09/2006
PATENT APPLICATION: **US/10/663,857A** TIME: 13:06:10

Input Set : **N:\efs\10663857a_efs\PX018432US_SEQ_1stOA_Amended.txt**
Output Set: **N:\CRF4\11092006\J663857A.raw**

```
193 <400> SEQUENCE: 2
194 Met Val Ser Lys Leu Ser Gln Leu Gln Thr Glu Leu Leu Ala Ala Leu
195   1               5                  10                  15
197 Leu Glu Ser Gly Leu Ser Lys Glu Ala Leu Ile Gln Ala Leu Gly Glu
198              20                  25                  30
200 Pro Gly Pro Tyr Leu Leu Ala Gly Glu Gly Pro Leu Asp Lys Gly Glu
201          35                  40                  45
203 Ser Cys Gly Gly Gly Arg Gly Glu Leu Ala Glu Leu Pro Asn Gly Leu
204      50                  55                  60
206 Gly Glu Thr Arg Gly Ser Glu Asp Glu Thr Asp Asp Asp Gly Glu Asp
207  65                  70                  75                  80
209 Phe Thr Pro Pro Ile Leu Lys Glu Leu Glu Asn Leu Ser Pro Glu Glu
210                  85                  90                  95
212 Ala Ala His Gln Lys Ala Val Val Glu Thr Leu Leu Gln Glu Asp Pro
213              100                 105                 110
215 Trp Arg Val Ala Lys Met Val Lys Ser Tyr Leu Gln Gln His Asn Ile
216          115                 120                 125
218 Pro Gln Arg Glu Val Val Asp Thr Thr Gly Leu Asn Gln Ser His Leu
219      130                 135                 140
221 Ser Gln His Leu Asn Lys Gly Thr Pro Met Lys Thr Gln Lys Arg Ala
222 145                 150                 155                 160
224 Ala Leu Tyr Thr Trp Tyr Val Arg Lys Gln Arg Glu Val Ala Gln Gln
225                 165                 170                 175
227 Phe Thr His Ala Gly Gln Gly Gly Leu Ile Glu Glu Pro Thr Gly Asp
228              180                 185                 190
230 Glu Leu Pro Thr Lys Lys Gly Arg Arg Asn Arg Phe Lys Trp Gly Pro
231          195                 200                 205
233 Ala Ser Gln Gln Ile Leu Phe Gln Ala Tyr Glu Arg Gln Lys Asn Pro
234      210                 215                 220
236 Ser Lys Glu Glu Arg Glu Thr Leu Val Glu Glu Cys Asn Arg Ala Glu
237 225                 230                 235                 240
239 Cys Ile Gln Arg Gly Val Ser Pro Ser Gln Ala Gln Gly Leu Gly Ser
240                 245                 250                 255
242 Asn Leu Val Thr Glu Val Arg Val Tyr Asn Trp Phe Ala Asn Arg Arg
243              260                 265                 270
245 Lys Glu Glu Ala Phe Arg His Lys Leu Ala Met Asp Thr Tyr Ser Gly
246          275                 280                 285
248 Pro Pro Pro Gly Pro Gly Pro Gly Pro Ala Leu Pro Ala His Ser Ser
249      290                 295                 300
251 Pro Gly Leu Pro Pro Pro Ala Leu Ser Pro Ser Lys Val His Gly Val
252 305                 310                 315                 320
254 Arg Tyr Gly Gln Pro Ala Thr Ser Glu Thr Ala Glu Val Pro Ser Ser
255                 325                 330                 335
257 Ser Gly Gly Pro Leu Val Thr Val Ser Thr Pro Leu His Gln Val Ser
258              340                 345                 350
260 Pro Thr Gly Leu Glu Pro Ser His Ser Leu Leu Ser Thr Glu Ala Lys
261          355                 360                 365
263 Leu Val Ser Ala Ala Gly Gly Pro Leu Pro Pro Val Ser Thr Leu Thr
264      370                 375                 380
```

RAW SEQUENCE LISTING DATE: 11/09/2006
PATENT APPLICATION: US/10/663,857A TIME: 13:06:10

Input Set : N:\efs\10663857a_efs\PX018432US_SEQ_1stOA_Amended.txt
Output Set: N:\CRF4\11092006\J663857A.raw

```
266 Ala Leu His Ser Leu Glu Gln Thr Ser Pro Gly Leu Asn Gln Gln Pro
267 385                 390                 395                 400
269 Gln Asn Leu Ile Met Ala Ser Leu Pro Gly Val Met Thr Ile Gly Pro
270                 405                 410                 415
272 Gly Glu Pro Ala Ser Leu Gly Pro Thr Phe Thr Asn Thr Gly Ala Ser
273             420                 425                 430
275 Thr Leu Val Ile Gly Leu Ala Ser Thr Gln Ala Gln Ser Val Pro Val
276         435                 440                 445
278 Ile Asn Ser Met Gly Ser Ser Leu Thr Thr Leu Gln Pro Val Gln Phe
279     450                 455                 460
281 Ser Gln Pro Leu His Pro Ser Tyr Gln Gln Pro Leu Met Pro Pro Val
282 465                 470                 475                 480
284 Gln Ser His Val Thr Gln Asn Pro Phe Met Ala Thr Met Ala Gln Leu
285                 485                 490                 495
287 Gln Ser Pro His Ala Leu Tyr Ser His Lys Pro Glu Val Ala Gln Tyr
288             500                 505                 510
290 Thr His Thr Gly Leu Leu Pro Gln Thr Met Leu Ile Thr Asp Thr Thr
291         515                 520                 525
293 Asn Leu Ser Ala Leu Ala Ser Leu Thr Pro Thr Lys Gln Val Phe Thr
294     530                 535                 540
296 Ser Asp Thr Glu Ala Ser Ser Glu Ser Gly Leu His Thr Pro Ala Ser
297 545                 550                 555                 560
299 Gln Ala Thr Thr Leu His Ile Pro Ser Gln Asp Pro Ala Gly Ile Gln
300                 565                 570                 575
302 His Leu Gln Pro Ala His Arg Leu Ser Ala Ser Pro Thr Val Ser Ser
303             580                 585                 590
305 Ser Ser Leu Val Leu Tyr Gln Ser Ser Asp Ser Ser Asn Gly Gln Ser
306         595                 600                 605
308 His Leu Leu Pro Ser Asn His Ser Val Ile Glu Thr Phe Ile Ser Thr
309     610                 615                 620
311 Gln Met Ala Ser Ser Ser Gln
312 625                 630
315 <210> SEQ ID NO: 3
316 <211> LENGTH: 93
317 <212> TYPE: DNA
318 <213> ORGANISM: Homo sapiens
320 <400> SEQUENCE: 3
321 gtaaggtcca ggcctgctgg ccctcccttg gcctgtgaca gagcccctca cccccacatc        60
323 ccccgggctc aggaggctgc tctgctcccc cag                                    93
326 <210> SEQ ID NO: 4
327 <211> LENGTH: 41
328 <212> TYPE: DNA
329 <213> ORGANISM: Artificial Sequence
331 <220> FEATURE:
332 <223> OTHER INFORMATION: sense primer for amplifying promoter of MODY3 gene
335 <400> SEQUENCE: 4
336 taatacgact cactataggg tggccgtgag catcctctgc c                             41
339 <210> SEQ ID NO: 5
340 <211> LENGTH: 39
```

VERIFICATION SUMMARY DATE: 11/09/2006
PATENT APPLICATION: **US/10/663,857A** TIME: 13:06:11

Input Set : **N:\efs\10663857a_efs\PX018432US_SEQ_1stOA_Amended.txt**
Output Set: **N:\CRF4\11092006\J663857A.raw**

L:8 M:270 C: Current Application Number differs, Replaced Current Application Number